Optimality and equilibria in stochastic games

Optimaliteit en Evenwichten in Stochastische Spelen

Samenvatting

Dit proefschrift handelt over tweepersoons stochastische spelen met eindige toestandsruimte en eindige aktieruimten. Dit zijn niet-coöperatieve dynamische spelen waarbij het spelproces zich, via een aftelbaar oneindig aantal tijdstippen, afspeelt over een eindig aantal toestanden. Ieder van de toestanden kan als begintoestand dienen voor het spelproces. Op ieder tijdstip bevindt het spelproces zich in precies één van die toestanden en moeten twee personen, speler 1 en speler 2 genoemd, gelijktijdig en onafhankelijk van elkaar, elk een aktie kiezen uit een bij die toestand behorende eindige aktieverzameling. Die toestand en de gekozen akties bepalen voor beide spelers een uitbetaling alsook een kansvector volgens welke het spelproces zich naar een volgende toestand verplaatst. Doel van elk van de spelers is het om zo veel mogelijk te verdienen en dat zonder de mogelijkheid te hebben om bindende afspraken te maken.

Op ieder tijdstip kennen beide spelers niet alleen de toestandsruimte, de aktieruimten en de bijbehorende uitbetalings- en overgangsfunktie, maar ook kennen zij de hele historie op dat tijdstip. Hierbij is de historie op een tijdstip de reeks van toestanden en aktiekeuzen (van beide spelers) volgens welke het spelproces zich tot dan toe heeft afgespeeld. Gebruikmakend van deze informatie hanteren de spelers strategieën, speelplannen, om het spelproces te sturen. Een strategie van een speler legt precies vast voor ieder tijdstip, voor iedere toestand en voor iedere historie, welke aktie die speler moet kiezen. De meest complexe strategieën zijn historie-afhankelijke strategieën. Daaronder worden vooral strategieën verstaan waarbij de te nemen aktiekeuzen afhankelijk zijn van, door de tegenstander, in het verleden gekozen akties. De minst complexe strategieën zijn stationaire strategieën, waarbij de voorgeschreven aktiekeuzen noch van het tijdstip, noch van de historie, maar enkel van de toestand afhangen.

De spelers proberen ieder hun verwachte opbrengst te maximaliseren. Hierbij wordt de verwachte opbrengst voor een speler vastgelegd door de manier waarop die speler de oneindige stroom van uitbetalingen interpreteert. Om die oneindige stromen van uitbetalingen te interpreteren gebruiken de spelers een evaluatiecriterium. We gaan ervan uit dat beide spelers hetzelfde criterium gebruiken. In dit proefschrift worden stochastische spelen bestudeerd voor drie verschillende evaluatiecriteria, te weten het β-verdisconteerde opbrengstencriterium, het limiet-gemiddelde opbrengstencriterium en het totale opbrengstencriterium.

In het algemeen zullen de spelers verschillende preferenties hebben over de toestanden en de daarin, onafhankelijk van elkaar, te kiezen aktieparen. Een interessante klasse van stochastische spelen wordt gevormd door de zogenaamde nulsom stochastische spelen waarbij, voor iedere toestand en voor ieder paar akties dat daar gekozen kan worden, geldt dat speler 2 datgene betaalt wat speler 1 krijgt. In dergelijke spelen hebben de spelers strikt tegengestelde belangen. De centrale vraag in nulsom stochastische spelen is de vraag of er (voor iedere starttoestand) een bepaald bedrag v bestaat, waarvoor geldt dat speler 1 een strategie heeft waarmee hij ervoor kan zorgen dat zijn verwachte opbrengst minstens v zal zijn en waarvoor gelijktijdig geldt dat speler 2 een strategie heeft waarmee die ervoor kan zorgen dat de verwachte opbrengst van speler 1 hoogstens v zal zijn. Mogelijk kunnen de spelers zich v slechts op ϵ na garanderen ($\epsilon>0$). Zo'n bedrag v wordt dan de waarde (voor die starttoestand) genoemd en de bijbehorende strategieën heten (ϵ-)optimale strategieën. Hierbij zijn waarde en ϵ-optimale strategieën natuurlijk afhankelijk van het, door beide spelers gehanteerde, evaluatiecriterium.

In niet-nulsom stochastische spelen hebben de begrippen 'waarde' en 'ϵ-optimale strategieën' weinig betekenis. Daar vormen ϵ-evenwichten het gebruikelijke oplossingsconcept. Een ϵ-evenwicht is een paar strategieën met de eigenschap dat geen der spelers meer dan ϵ kan verdienen door eenzijdig van zijn strategie af te wijken.

In dit proefschrift worden zowel nulsom stochastische spelen als niet-nulsom stochastische spelen bestudeerd, en wel voor ieder van de drie hierboven genoemde evaluatiecriteria.

Voor stochastische spelen ten aanzien van het β-verdisconteerde opbrengstencriterium is het welbekend dat in het nulsom geval de waarde en stationaire optimale strategieën voor beide spelers altijd bestaan; in het niet-nulsom geval bestaan er altijd stationaire evenwichten.

Ten aanzien van het limiet-gemiddelde opbrengstencriterium bestaat in het nulsom geval de waarde, maar voor beide spelers zullen historie-afhankelijke strategieën in het algemeen onontbeerlijk zijn om ϵ-optimaal te kunnen spelen; in het niet-nulsom geval is de existentie van ϵ-evenwichten het belangrijkste onopgeloste probleem.

Voor wat het totale opbrengstencriterium betreft zijn er slechts oplossingen bekend voor speciale gestructureerde klassen van stochastische spelen. Ook voor dit laatste criterium zijn historie-afhankelijke strategieën onmisbaar om tot een oplossing te komen.

De indeling van dit proefschrift is als volgt.

In hoofdstuk 1 worden stochastische spelen formeel geïntroduceerd, alsook de erbij behorende oplossingsconcepten. Verder wordt een beknopt overzicht gegeven van historische resultaten die voor de onderwerpen in dit proefschrift van belang zijn. Voorts worden enkele elementaire resultaten afgeleid. Een en ander wordt aan de hand van een aantal voorbeelden toegelicht.

In hoofdstuk 2 wordt aangetoond dat er voor ieder niet-nulsom stochastisch spel een niet-lege verzameling begintoestanden is, waarvoor een 'bijna stationair' limiet-gemiddeld ϵ-evenwicht bestaat. Met een bijna stationair limiet-gemiddeld ϵ-evenwicht bedoelen we een ϵ-evenwicht, ten aanzien van het limiet-gemiddelde opbrengsten criterium, bestaande uit strategieën van het

type: gebruik een bepaalde stationaire strategie tenzij het zeer waarschijnlijk is dat je tegenstander afwijkt van een eveneens bepaalde strategie; in dat laatste geval, begin onmiddellijk je tegenstander af te straffen door optimaal te spelen in het gerelateerde nulsom spel dat gegeven wordt door de uitbetalingen aan die tegenstander.
Voor nulsom stochastische spelen laten we zien dat er voor iedere speler een niet-lege verzameling begintoestanden is, van waaruit die speler een stationair limiet-gemiddeld optimale strategie heeft. We geven voldoende voorwaarden onder welke beide spelers stationaire limiet-gemiddeld ϵ-optimale strategieën hebben voor alle begintoestanden met maximale, dan wel minimale, limiet-gemiddelde waarde. Een voorbeeld laat zien dat er begintoestanden kunnen zijn waarvoor geen der spelers een stationaire limiet-gemiddeld ϵ-optimale strategie heeft.

In hoofdstuk 3 worden de resultaten uit hoofdstuk 2 voor niet-nulsom stochastische spelen uigebreid door voldoende voorwaarden te geven voor het bestaan van een bijna stationair limiet-gemiddeld ϵ-evenwicht voor alle begintoestanden. Deze voldoende voorwaarden worden geformuleerd aan de hand van eigenschappen van een convergente rij van stationaire β-verdisconteerde evenwichten (waarbij β naar 1 gaat).

In hoofdstuk 4 bestuderen we implicaties van de resultaten uit de hoofdstukken 2 en 3 voor een aantal deelklassen van stochastische spelen: stochastische spelen met voor ieder paar stationaire strategieën slechts één ergodische klasse, stochastische spelen met toestandsonafhankelijke overgangen, herhaalde spelen met absorberende toestanden.

Hoofdstuk 5 is gewijd aan het totale opbrengstencriterium. We laten zien dat, zelfs wanneer de limiet-gemiddelde waarde 0 is voor iedere toestand, de totale waarde niet hoeft te bestaan. Ook tonen we aan dat zelfs onder de sterkere conditie dat de limiet-gemiddelde waarde 0 is en dat beide spelers stationaire limiet-gemiddeld optimale strategieën bezitten, historie-afhankelijke strategieën onmisbaar kunnen zijn om totaal ϵ-optimaal te spelen. Onder deze sterkere voorwaarde geven we ook karakteriseringen voor het bestaan van de totale waarde en stationaire totaal optimale strategieën voor beide spelers. Tevens vergelijken we het totale opbrengstencriterium met zowel het β-verdisconteerde opbrengstencriterium als het limiet-gemiddelde opbrengstencriterium.

In hoofdstuk 6 worden karakteriseringen gegeven voor het bestaan van stationaire evenwichten (resp. (ϵ-)optimale strategieën) in niet-nulsom (resp. nulsom) stochastische spelen, aan de hand van oplossingen voor gerelateerde mathematische programmas. Dit wordt gedaan voor elk van de 3 evaluatie-criteria, waarbij we voor het totale opbrengstencriterium veronderstellen dat de limiet-gemiddelde opbrengst 0 is voor alle paren van stationaire strategieën.